

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 16, 2003, 15:50:54 : Search time 2492.14 Seconds  
(without alignments)  
16465.781 Million cell updates/sec

Title: us-09-497-967-44

Perfect score: 1410

Sequence: 1 atgaataataatttagt.....cttattattattatgatga 1410

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_hgt.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pi.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: gb\_vt.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_other.\*

33: em\_htg\_mus.\*

34: em\_htg\_pin.\*

35: em\_htg\_rod.\*

36: em\_htg\_mam.\*

37: em\_htg\_vrt.\*

38: em\_sy.\*

39: em\_htgo\_hum.\*

40: em\_htgo\_mus.\*

41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1410	100.0	3026	3	AF324424	Ichthyoph
2	259.8	18.4	1520	3	AF405431	Ichthyoph
3	258	18.3	2486	3	AF140273	Ichthyoph
4	214.6	15.2	1249	3	ICYMANT	Ichthyoph
C 5	96	6.8	175748	2	AC120669	Rattus no
C 6	96	6.8	180903	2	AC125859	Rattus no
C 7	95.4	6.8	186935	2	AC022322	Rattus no
C 8	92.6	6.6	179553	2	AC024253	Homo sapi
C 9	90.8	6.4	131274	2	AC098669	Homo sapi
C 10	90.8	6.4	176822	2	AC098669	Homo sapi
C 11	85.6	6.1	261604	2	AC099174	Rattus no
C 12	85.4	6.1	127354	2	AC119819	Rattus no
C 13	85.4	6.1	155019	2	AC117014	Mus muscu
C 14	81	5.7	35793	5	AY016024	Rattus no
C 15	81	5.7	84472	2	AC098684	Takifugu
C 16	80.2	5.7	132449	9	AL365272	Takifugu
C 17	77.2	5.5	22398	5	FRU271723	Human DNA
C 18	77	5.5	85786	9	AL162582	Fugu rubr
C 19	76.6	5.4	10115	5	AF397467	Human DNA
C 20	75.8	5.4	183413	2	AC131200	Ictalurus
C 21	75.6	5.4	170985	2	AC096032	Rattus no
C 22	74.6	5.3	204259	2	AC110817	Rattus no
C 23	74	5.2	172307	2	AC044842	Rattus no
C 24	72.4	5.1	189461	2	AC119627	Homo sapi
C 25	72.4	5.1	191841	2	AC121374	Rattus no
C 26	72	5.1	134558	2	AC125757	Rattus no
C 27	72	5.1	180668	2	AC020857	Mus muscu
C 28	69.8	5.0	666	8	AF413050S2	Ze mays
C 29	69.2	4.9	172853	9	AC084361	Homo sapi
C 30	68.2	4.8	71553	2	AC096997	Takifugu
C 31	68	4.8	154141	2	AC121042	Rattus no
C 32	68	4.8	162269	9	AC000097	Homo sapi
C 33	67.8	4.8	184830	2	AC113220	Rattus no
C 34	67.2	4.8	185994	2	AC002042	Rattus no
C 35	67.2	4.8	300994	2	AC129317	Homo sapi
C 36	66.2	4.7	131346	2	AC119558	Mus muscu
C 37	66	4.7	158615	2	AC117835	Rattus no
C 38	65.8	4.7	101534	2	AC108576	Rattus no
C 39	65.8	4.7	298283	3	AE003782	Rattus no
C 40	64.8	4.6	177035	2	AC099443	Drosophill
C 41	64.4	4.6	186558	2	AC079031	Rattus no
C 42	64.4	4.6	325808	2	AC121787	Homo sapi
C 43	64.2	4.6	187617	2	AC118993	Mus muscu
C 44	64	4.5	100726	2	AC116961	Rattus no
C 45	63.8	4.5	177435	2	AC128321	Dictyoste

ALIGNMENTS

RESULT 1  
AF324424  
LOCUS  
DEFINITION Ichthyophthirius multifiliis 3026 bp DNA linear INV 27-FEB-2002  
(IAGS2A) gene, complete cds.  
ACCESSION AF324424  
VERSION AF324424.1 GI:12698726  
KEYWORDS  
SOURCE Ichthyophthirius multifiliis.  
ORGANISM Ichthyophthirius multifiliis.  
Eukaryota; Alveolata; Ciliophora; Oligohymenophorea;  
Hymenostomatida; Ophryoglenina; Ichthyophthirius.  
REFERENCE  
1. (bases 1 to 3026)  
Lin,Y., Lin,T.L., Wang,C.C., Wang,X., Stieger,K., Klopfleisch,R.  
and Clark,T.G.

TITLE Variation in primary sequence and tandem repeat copy number among  
i-antigens of Ichthyophthirius multifiliis  
JOURNAL Mol. Biochem. Parasitol. 120 (1), 93-106 (2002)  
MEDLINE 21839613  
PUBMED 11849709  
REFERENCE 2 (bases 1 to 3026)  
AUTHORS Lin, Y., Lin, T.-L. and Clark, T. G.  
TITLE Direct Submission  
JOURNAL Submitted (27-NOV-2000) Microbiology and Immunology, NYSCVM Cornell  
University, Ithaca, NY 14853, USA  
FEATURES  
source Location/Qualifiers  
1..3026  
/organism="Ichthyophthirius multifiliis"  
/macronuclear  
/serotype="D"  
/isolate="G5"  
/db\_xref="taxon:5932"  
1285..2817  
/gene="IAG52A"  
1285..2817  
/gene="IAG52A"  
/product="immobilization antigen isoform"  
1330..2736  
/gene="IAG52A"  
/note="52 kDa protein"  
/codon\_start=1  
/transl\_table=6  
/product="immobilization antigen isoform"  
/protein\_id="AAK01661.1"  
/db\_xref="GI:12698727"  
/translation="MKNTLVLLIISLFINOIKSANCVPVGTETNAGOVDLGTPANC  
VNCQNFYNNAAEPVPGASTCTPCPKQKDAQNPPTANLVTCNVKPCPACTAIA  
GGATDYAAIITEVCNCRINFENAPNENAGASTCTACPNVRVGGALTAGNAATVIAQ  
CNVACPTGLDDGVITDYVSFTECVKRLNFYNGNNGNTFPNPKSGCQCPKAIK  
PANVAQATLGNDAITPAQCNVACPDGTISAAGVNNVNAONTCTNACAPFNFNNAFNP  
NPNCSNCLCPKANDYGAETAGGAATLAKOCNIAACPDGTAIASGATNVVLIQTECLN  
CAANFYDDNPNQGSRRCKACPKANKVOGAVATAGGTATLIAQCALECPAGTVLTDGT  
TSYKOAASECVKCANFYTKQTDWVAGIDFCTSCNKKLTSGAEANLPESAKKNIQC  
DFANFLSISLILLLISYLL"

BASE COUNT 1177 a 316 c 355 g 1178 t  
ORIGIN  
Query Match 100.0%; Score 1410; DB 3; Length 3026;  
Best Local Similarity 100.0%; Pred. No. 4.1e-237;  
Matches 1410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAATAATTTTAGTAAATATGATTTTCATTTATTCATTTATCAATTAATCT 60  
DB 1330 ATGAAAAATAATTTTAGTAAATATGATTTTCATTTATTCATTTATCAATTAATCT 1389  
QY 61 GCTAATTCCTGTTGGAACCTGAACCTAACACACGCGGATAAGTTGATGCTAGGAAC 120  
DB 1390 GCTAATTCCTGTTGGAACCTGAACCTAACACACGCGGATAAGTTGATGCTAGGAAC 1449  
QY 121 CTGCAAAATGTTGTTAAATGTTAGAAAACTTTTATTATAAATGCTGCTTTGCTT 180  
DB 1450 CTGCAAAATGTTGTTAAATGTTAGAAAACTTTTATTATAAATGCTGCTTTGCTT 1509  
QY 181 CTGGTGCTAGTACGTGACCTTGCCATAAAAAAGATGCTGCTTAACCAAT 240  
DB 1510 CTGGTGCTAGTACGTGACCTTGCCATAAAAAAGATGCTGCTTAACCAAT 1569  
QY 241 CCACCTGCTACGCTAAATTTAGTACATATGTAACGTTAAATGCCCTGCTGGTACCGCA 300  
DB 1570 CCACCTGCTACGCTAAATTTAGTACATATGTAACGTTAAATGCCCTGCTGGTACCGCA 1629  
QY 301 ATTGCAGGTGGAGCAACAGATTATCGACAATAATPCACAGAAATGTGTTAATTTAGAAAT 360  
DB 1630 ATTGCAGGTGGAGCAACAGATTATGAGCAATAATPCACAGAAATGTGTTAATTTAGAAAT 1689  
QY 361 AATTTTATAAAGAAATGCTCAATTTTAAATGAGGTGCTAGTACATGACAGCTTGT 420  
DB 1690 AATTTTATAAAGAAATGCTCAATTTTAAATGAGGTGCTAGTACATGACAGCTTGT 1749

QY 421 CCGGTAACACAGAGTTGGTGGTGCAATTTGACTGCTGTTAATGCCCTACCATAGTCGCATAA 480  
DB 1750 CCGGTAACACAGAGTTGGTGGTGCAATTTGACTGCTGTTAATGCCCTACCATAGTCGCATAA 1809  
QY 481 TGTAAAGCTGCGCATGCTCTACTGGTACTGCACTTGATGAGGAGTAACACTACTGATTGTT 540  
DB 1810 TGTAAAGCTGCGCATGCTCTACTGGTACTGCACTTGATGAGGAGTAACACTACTGATTGTT 1869  
QY 541 AGATCAATTCACAGAAATGTTAAATGTAAGTAAATGTAAGTAAATGTAAGTAAATGTAAGT 600  
DB 1870 AGATCAATTCACAGAAATGTTAAATGTAAGTAAATGTAAGTAAATGTAAGTAAATGTAAGT 1929  
QY 601 AATACTCCTTTCAATCCAGGTAAAGTAAATGACACACCTTGTCCGGCAATTAACCTGCT 660  
DB 1930 AATACTCCTTTCAATCCAGGTAAAGTAAATGACACACCTTGTCCGGCAATTAACCTGCT 1989  
QY 661 AATGTTGCTTAAAGCTACTTTAGGTAATGATGATCAATAACCGCATATGTAACGTTGCA 720  
DB 1990 AATGTTGCTTAAAGCTACTTTAGGTAATGATGATCAATAACCGCATATGTAACGTTGCA 2049  
QY 721 TGGCCTGATGCTACTATAAGTCTGCTGAGTAAATTAATTTGGTAGCACAAAACACTGAA 780  
DB 2050 TGGCCTGATGCTACTATAAGTCTGCTGAGTAAATTAATTTGGTAGCACAAAACACTGAA 2109  
QY 781 TGTACTAATTTGCTCTCTTAACCTTTTACAATAATAATGCTCCTTAATTTCAATCCAGGTAAT 840  
DB 2110 TGTACTAATTTGCTCTCTTAACCTTTTACAATAATAATGCTCCTTAATTTCAATCCAGGTAAT 2169  
QY 841 AGTACATGCTACTCTGCCAGCAATAAAGATTATGTTGCTGAAGCCACTGCAGGTTGCT 900  
DB 2170 AGTACATGCTACTCTGCCAGCAATAAAGATTATGTTGCTGAAGCCACTGCAGGTTGCT 2229  
QY 901 GCCCTACTTTAGCCAAATTAATTAATTTGATGCTGCTGATGCTGATGCTGATGCTGATGCT 960  
DB 2230 GCCCTACTTTAGCCAAATTAATTAATTTGATGCTGCTGATGCTGATGCTGATGCTGATGCT 2289  
QY 961 GGAGCAACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1020  
DB 2290 GGAGCAACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2349  
QY 1021 TTTGATGTAATAATTTCTAGGAGGAAAGTAGTAGTGAAGCAATGTCACGCAATAATAA 1080  
DB 2350 TTTGATGTAATAATTTCTAGGAGGAAAGTAGTAGTGAAGCAATGTCACGCAATAATAA 2409  
QY 1081 GTTAAAGCGCTGTAGCAACTGCAAGTGGTACTGCTACTTTAATTCGATAATGTCCTTT 1140  
DB 2410 GTTAAAGCGCTGTAGCAACTGCAAGTGGTACTGCTACTTTAATTCGATAATGTCCTTT 2469  
QY 1141 GAATGCCCTGCTGCTACTGCTACTCACCAGGAGGAAAGTAGTAGTGAAGCAATAATAA 1200  
DB 2470 GAATGCCCTGCTGCTACTGCTACTCACCAGGAGGAAAGTAGTAGTGAAGCAATAATAA 2529  
QY 1201 TCTGAATGTTAAATGTTGCTGCAACTTTTATTAATTAATTAATTAATTAATTAATTAATTA 1260  
DB 2530 TCTGAATGTTAAATGTTGCTGCAACTTTTATTAATTAATTAATTAATTAATTAATTAATTA 2589  
QY 1261 GGTATGATGATGCTACTGCTACTGTTAAATAAATAATTAATTTCTGGCGCTGAAGCTAATTTA 1320  
DB 2590 GGTATGATGATGCTACTGCTACTGTTAAATAAATAATTAATTTCTGGCGCTGAAGCTAATTTA 2649  
QY 1321 CCTGAATGCTGTAATAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1380  
DB 2650 CCTGAATGCTGTAATAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2709  
QY 1381 TTATTGATTTCTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 1410  
DB 2710 TTATTGATTTCTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 2739

RESULT 2  
AF405431  
LOCUS AF405431 1520 bp mRNA linear INV 27-FEB-2002





Db 1613 ATACATGACTGAAATGTACTAAAAAATTAATCTCTGCTGCCACAGCTAAAGTATATGCTG 1672

Qy 1328 CTGCTAAAAAATATATAATG-----TGATTTTCGCTAATTTTATCAATTTCCCT 1378

Db 1673 AAGTACTCAAAAGTATAATGCGCCTCCACTACTTTTCGCTAAATTTTATCGATTTCCT 1732

Qy 1379 TATTATTGATTCTCTTATTATTATTATGATGA 1410

Db 1733 TATTATTATTCTTCTCTATTATTATGATGA 1764

RESULT 4

ICYMANT

LOCUS

DEFINITION Ichthyophthirius multifiliis 1249 bp mRNA linear INV 18-SEP-1998

ACCSSION M92907

VERSION M92907.1 GI:3628568

KEYWORDS

SOURCE

ORGANISM

Ichthyophthirius multifiliis.

Eukaryota: Alveolata: Ciliophora: Oligohymenophorea;

Hymenostomida; Ophryoglenina; Ichthyophthirius.

1. (bases 1 to 1249)

Lin,T.L. and Dickerson,H.W.

Purification and partial characterization of immobilization

antigens from Ichthyophthirius multifiliis

J. Protozool. 39 (4), 457-463 (1992)

93020590

1383510

2. (bases 1 to 1249)

Clark,T.G., McGraw,R.A. and Dickerson,H.W.

Developmental expression of surface antigen genes in the parasitic

ciliate Ichthyophthirius multifiliis

Proc. Natl. Acad. Sci. U.S.A. 89 (14), 6363-6367 (1992)

92335298

1631132

3. (bases 1 to 1249)

Clark,T.

Direct Submission

Submitted (06-JUN-1992) Microbiology and Immunology, Cornell

University, Ithaca, NY 14853, USA

4. (bases 1 to 1249)

Clark,T.

Direct Submission

Submitted (18-SEP-1998) Microbiology and Immunology, Cornell

University, Ithaca, NY 14853, USA

Sequence update by submitter

On Sep 18, 1998 this sequence version replaced gi:159289.

Location/Qualifiers

1. .1249

/organism="Ichthyophthirius multifiliis"

/strain="Georgia"

/db\_xref="taxon:5932"

/dev\_stage="trophont"

<1. .1190

/note="48 kDa"

/citation=[2]

/citation=[1]

/codon\_start=3

/evidence=experimental

/transl\_table=6

/product="immobilization antigen precursor"

/db\_xref="GI:3628568"

/translation="AVPCPDGTQTQAGLTDVGADLGTVCNCRPNFYNGGAAQGBAN

RSQPFANNAARGICPCQINRVGSVYNAGDLATLQCTSTOCPTGTLDDGVDVDF

NSAQCKVKCPNFYNGGSPQSGPQVFAAGAAAAGVAAVTSQVPCQLNKNDSPA

TAGAQLNATQCSNOCPGTGLVLDGTVLVNTSATLVCVKRPNFYNGGSPQSGAPGV

QVFAAGAAAGVAAVTSQVPCQINKNDSPATAGAAANLATQCTSTOCPTGTLQDQVT

LVSNSQCSQCIANYFFNGNLEGRKSCLKCPVSKTTPAHAPGNATQATQCLITC

PAGTVLDDGTSTNFEVASATEBCTKCSAGFFASKTTGTAGTDCTCTCKLTSGATAVN

IHQ"

sig\_peptide

mat\_peptide

repeat\_region

BASE COUNT 348 a 243 c 257 g 401 t

ORIGIN

Query Match 15.2%; Score 214.6; DB 3; Length 1249;

Best Local Similarity 55.4%; Pred. No. 5.4e-28;

Matches 591; Conservative 0; Mismatches 374; Indels 102; Gaps 5;

Qy 344 GTGTTAATTGTAGAAATTAATTTTATTAATGAAATGCTCCAAATTTTAAATGAGGTGCTA 403

Db 112 GTGCTGCTTTAAGGGAAGCTAATGTAATTAACCTTTTCGACGCAAAATAATGCTGTAGAG 171

Qy 404 GTACATGCACAGCTTGTCCGGTAAACAGAGTTGGTGGTGCATTTGACTGTGTAATGCCG 463

Db 172 GTATATGTGTACCATGCCAAATTAACAGAGTAGGCTCTGTTACCAATGCAGGTGACTTAG 231

Qy 464 CTACCATAGTCGCATTAATGTAAAGTCGCATGCTCTACTGCTACTGCACPTTGATGGAG 523

Db 232 CTACTTTAGCCACATTAATGCAGTACTTAATGTCCTACTGGCACTGCACCTTGATGGAG 291

Qy 524 TAACTACTGATTATGTAGATCATTCACAGAAATGTTAAATGTAGACTTAACCTTTTACT 583

Db 292 TGACAGATGTTTTGTATAGATCAGCCGCAATATGTTTAAATGCAACCTAACTTTTACT 351

Qy 584 ATAATGGTAAATATGTTAATCTACTCCCTTCAATCCAGGTAAAAGTTAATGCACACACTTGC 643

Db 352 ATAATGGTGGTCTCTCTTAAGGTGAAGCTCCTGGCGCTTAAGTTTTCGCTGGTGGCTG 411

Qy 644 CGCAATTAACACCTGCT-----ATG 664

Db 412 CGCTGCAGGTGTGCTGCCGTTACTAGTTAATGTACTTGTCCCACTAAACAAAAAGC 471

Qy 665 TTGCTTAAAGTACTTTAGTAAATGCTGCTACAAATACCGCATANTGTAAGTTGATGCC 724

Db 472 ATTCTCTGCCACTGCAGCGCTTAAGCTAAATTTAGCCACATAATGTAACAATTAATGTC 531

Qy 725 CTGATGCTACTATAAGTGTCTGCTGGAGT---AAATAAATTTGGGTAGCACAAAACACTGAAT 781

Db 532 CTACTGGCACTGTACTTTGATGAGTGACACTTCTTTTAAATACATCAGCCACATTAAT 591

Qy 782 GTACTAATTTGCTCTCTAACTTTTACAATAATAATGCTCCTAAT-----826

Db 592 GTGTTAAATGCAGACCTAACTTTTACTAATAATGGTGGTTCCTCTTAAAGGTGAAGCTCC 651

Qy 827 -----TCAATCCAGGTAAATAGTACAT 847

Db 652 GCCTTAAAGTTTTGTGCTGGTGGCTGCCGCTGCAGGTGTGCTGCCGTTACTAGTTAAT 711

Qy 848 GCCTACTTTGCCAGCAAAATAAGATTATGGTGTCAAGCCACTGCAGGTGGTCCGGCTA 907

Db 712 GTGTACTTTGCCAAATTAACAAAAACGATTCTCCT---GCCACTGCAGGTGCCTAGCTA 768

Qy 908 CTTTAGCCAAATAATCTAATATTGTCATGCCCTGATGGTACTGCAATTTGCTAGTGGAGCAA 967

Db 769 ATTTAGCCACATAATGCAGTACTTAATGTCCAACCTGGCAATTCGAATTAAGACGGAGTGA 828

Qy 968 CTAATTATGTAATTTATTAACAGAAATGCTTAATTTGCTGCTACTACTTTTATTTTATG 1027

Db 829 CACTTGTTTTGTAGTAATTCATCCACATAATGTTCTTAAATGCAATTCGTAATTTACTTT 888

Qy 1028 GTAATAATTCTAGGCAAGAGTAGTAGATGCAACCATGTGCAGCAAAATAAGTTTAAAG 1087

Db 889 ATGGTAAATCTCGAAGCAGGTAAAGTTAATGTTTAAAGTGTCCAGTTAGTAAGTAAC 945

Qy 1088 GCCTGTAGCAACTGCAGGTGGTGTACTGCTACTTTTAATTCATAATGTGCCCTTTGAATGC 1147

Db 946 CTCACGACATGCTCCAGGTATATCTGCTACTTAGCCACATAATGTTTGACCACATGTC 1005  
QY 1148 CTGCTGGTACTGTACTACCCGATGGAACAACATCTACTTATATAAAGCAGCATCTGAAT 1207  
Db 1006 CTGCTGGTACAGTACTGTATGATGGAACATCAACTAAATTTGTAGCTCCGCCAACTGAT 1065  
QY 1208 GTGTTAAATGTGTCGCAACTTTTATCTACTACAAAATAAATGATTTGGGTAGCAGGTATTG 1267  
Db 1066 GTACTAAATGTTCTGCTGGCTTTTGTGCATCAAAAACAACTGGTTTTCACAGCAGGTACTG 1125  
QY 1268 ATACATGTACTAGTCTGATATAAAAAATTAACCTTCCTGGCCCTGAAGCT 1314  
Db 1126 ATACATGTACTGAATGACTACTAAAAAATTAACCTTCCTGGTCCACAGCT 1172

## RESULT 5

AC120669/c

LOCUS Rattus norvegicus clone CH230-22G10, \*\*\* SEQUENCING IN PROGRESS  
DEFINITION \*\*\*, 46 unordered pieces.

AC120669

AC120669.2 GI:21902861

VERSION HTG; HTGS\_PHASE1.

SOURCE Rattus norvegicus.

ORGANISM Rattus norvegicus.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1 (bases 1 to 175748)

## AUTHORS

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
Alsbrooks,S.L., Amarantunge,H.C., Are,J.R., Ayelle,M., Banks,T.,  
Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,  
Bouch,J., Bowle,S., Brieve,M., Brown,E., Brown,M., Bryant,N.P.,  
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Y., Chowdhry,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,  
Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,  
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,E., Mathiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,  
Miner,G., Minal,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
Moser,M., Neer,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nwokenwo,S., Oguh,M., Okwuonu,G.,  
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,  
Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,  
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
Williams,G., Williamson,A., Wlarczyk,R., Wooden,S., Worley,K.,  
Wu,C., Wu,Y., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G. and Gibbs,R.

Direct Submission

Unpublished

REFERENCE 2 (bases 1 to 175748)

## AUTHORS

Worley,K.C.

TITLE  
JOURNALREFERENCE  
AUTHORS  
TITLE  
JOURNAL

## COMMENT

Direct Submission  
Submitted (09-MAY-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 175748)  
Worley, K.C.  
Direct Submission  
Submitted (23-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Jul 18, 2002 this sequence version replaced gi:20514530.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GXLIT

Center clone name: CH230-22G10

----- Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 130722 bases at least Q40

Consensus quality: 138187 bases at least Q30

Consensus quality: 143539 bases at least Q20

-----

\* NOTE: Estimated insert size may differ from sequence length

\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 46 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

\* 1 1312: contig of 1312 bp in length

\* 1313 1412: gap of unknown length

\* 1413 2562: contig of 1150 bp in length

\* 2563 2662: gap of unknown length

\* 2663 3966: contig of 1304 bp in length

\* 3967 4066: gap of unknown length

\* 4067 5262: contig of 1196 bp in length

\* 5263 5363: gap of unknown length

\* 5363 6793: contig of 1431 bp in length

\* 6794 6893: gap of unknown length

\* 6894 8272: contig of 1379 bp in length

\* 8273 8372: gap of unknown length

\* 8373 9620: contig of 1248 bp in length

\* 9621 9721 11136: contig of 1416 bp in length

\* 9721 11137: gap of unknown length

\* 11137 12579: contig of 1343 bp in length

\* 12580 12679: gap of unknown length

\* 12680 14466: contig of 1787 bp in length

\* 14467 14566: gap of unknown length

\* 14567 16613: contig of 2047 bp in length

\* 16614 16713: gap of unknown length

\* 16714 18834: contig of 2121 bp in length

\* 18835 18934: gap of unknown length

\* 18935 20996: contig of 2062 bp in length

\* 20997 21096: gap of unknown length

\* 21097 22643: contig of 1547 bp in length

\* 22644 22743: gap of unknown length

\* 22744 25384: contig of 2641 bp in length

\* 25385 25484: gap of unknown length

\* 25485 27048: contig of 1564 bp in length

\* 27049 27149: gap of unknown length

\* 27149 29221: contig of 2073 bp in length

\* 29222 29321: gap of unknown length

\* 29322 30973: contig of 1652 bp in length

\* 30974 31073: gap of unknown length





Eukaryota; Metazoa; Chordata; Craniata; Vertebrata: Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

# REFERENCE AUTHORS

1 (bases 1 to 180903)  
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Alsbrooks, S.L., Amarantunge, H.C., Are, J.R., Ayele, M., Banks, T., Barbarella, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buha, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mel, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G., Oraguie, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I., Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usman, K., Vasquez, L., Vara, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wlezyk, R., Woodson, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G. and Gibbs, R.

# TITLE JOURNAL

REFERENCE  
Worley, K.C.  
Direct Submission  
Submitted (02-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 180903)  
Worley, K.C.  
Direct Submission  
Submitted (23-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

# REFERENCE AUTHORS

TITLE  
JOURNAL  
Submitted (23-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

# COMMENT

Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: GWUQ  
Center clone name: CH230-35H21  
----- Summary Statistics  
Sequencing vector: Plasmid;  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 117810 bases at least Q40  
Consensus quality: 124839 bases at least Q30  
Consensus quality: 130161 bases at least Q20

-----  
\* NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 69 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 1018: contig of 1018 bp in length  
\* 1019 1118: gap of unknown length  
\* 1119 2397: contig of 1279 bp in length  
\* 2398 2497: gap of unknown length  
\* 2498 3536: contig of 1039 bp in length  
\* 3537 3636: gap of unknown length  
\* 3637 5038: contig of 1402 bp in length  
\* 5039 5138: gap of unknown length  
\* 5139 6416: contig of 1278 bp in length  
\* 6417 6516: gap of unknown length  
\* 6517 7882: contig of 1266 bp in length  
\* 7883 9056: contig of 1174 bp in length  
\* 9057 9156: gap of unknown length  
\* 9157 10223: contig of 1067 bp in length  
\* 10224 10323: gap of unknown length  
\* 10324 11849: contig of 1526 bp in length  
\* 11850 11949: gap of unknown length  
\* 11950 13459: contig of 1510 bp in length  
\* 13460 13559: gap of unknown length  
\* 13560 15061: contig of 1502 bp in length  
\* 15062 15161: gap of unknown length  
\* 15162 16401: contig of 1240 bp in length  
\* 16402 16501: gap of unknown length  
\* 16502 17501: contig of 1001 bp in length  
\* 17503 17602: gap of unknown length  
\* 17603 18331: contig of 1229 bp in length  
\* 18332 18931: gap of unknown length  
\* 18932 20133: contig of 1202 bp in length  
\* 20134 20233: gap of unknown length  
\* 20234 21645: contig of 1412 bp in length  
\* 21646 21745: gap of unknown length  
\* 21746 23608: contig of 1863 bp in length  
\* 23609 23708: gap of unknown length  
\* 23709 25221: contig of 1513 bp in length  
\* 25222 25321: gap of unknown length  
\* 25322 26764: contig of 1443 bp in length  
\* 26765 26864: gap of unknown length  
\* 26865 28249: contig of 1385 bp in length  
\* 28250 28349: gap of unknown length  
\* 28350 30014: contig of 1665 bp in length  
\* 30015 30114: gap of unknown length  
\* 30115 32030: contig of 1916 bp in length  
\* 32031 32130: gap of unknown length  
\* 32131 35243: contig of 3113 bp in length  
\* 35244 35343: gap of unknown length  
\* 35344 37525: contig of 2182 bp in length  
\* 37526 37625: gap of unknown length  
\* 37626 39405: contig of 1780 bp in length  
\* 39406 41253: gap of unknown length  
\* 41254 41353: gap of unknown length  
\* 41354 43191: contig of 1838 bp in length  
\* 43192 43291: gap of unknown length  
\* 43292 44918: contig of 1627 bp in length  
\* 44919 45018: gap of unknown length  
\* 45019 46790: contig of 1772 bp in length  
\* 46791 46890: gap of unknown length  
\* 46891 48766: contig of 1876 bp in length  
\* 48767 48866: gap of unknown length  
\* 48867 50304: contig of 1438 bp in length  
\* 50305 50404: gap of unknown length



[illegible]

Center code: WUGSC  
Web site: <http://genome.wustl.edu/gsc/index.shtml>  
----- Project Information -----  
Center project name: H.NH0317M22  
----- Summary Statistics -----  
Sequencing vector: M13; 100%  
Sequencing vector: plasmid; 0%  
Chemistry: Dye-terminator ET; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 181940 bases at least Q40  
Consensus quality: 183253 bases at least Q30  
Consensus quality: 183774 bases at least Q20  
Insert size: 186000; agarose-fp  
Insert size: 185435; sum-of-contigs  
Quality coverage: 6.94 in Q20 bases;  
Quality coverage: 7.00 in Q20 bases; sum-of-contigs  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 16 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\*  
\* 1 1363: contig of 1363 bp in length  
\* 1364 1463: gap of unknown length  
\* 1464 3253: contig of 1790 bp in length  
\* 3254 3353: gap of unknown length  
\* 3354 4598: contig of 1245 bp in length  
\* 4599 4698: gap of unknown length  
\* 4699 5833: contig of 1135 bp in length  
\* 5834 5933: gap of unknown length  
\* 5934 8647: gap of 2713 bp in length  
\* 8647 11793: contig of 3046 bp in length  
\* 11793 14813: contig of 2921 bp in length  
\* 14814 14913: gap of unknown length  
\* 14914 19837: contig of 4924 bp in length  
\* 19838 24772: gap of unknown length  
\* 24773 30965: contig of 4835 bp in length  
\* 30966 31065: gap of unknown length  
\* 31066 40841: contig of 6093 bp in length  
\* 40842 40941: contig of 9776 bp in length  
\* 40942 52869: gap of unknown length  
\* 52870 52969: contig of 11928 bp in length  
\* 52970 63635: gap of unknown length  
\* 63636 90222: contig of 10666 bp in length  
\* 90223 90323: gap of unknown length  
\* 90324 135921: contig of 26487 bp in length  
\* 135922 136021: gap of unknown length  
\* 136022 186935: contig of 50914 bp in length.  
FEATURES  
source  
1. .186935  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/clone="RP11-317M22"  
1. 1363  
/note="assembly\_name:Contig7"  
1464. .3253  
/note="assembly\_name:Contig8"  
3354. .4598  
/note="assembly\_name:Contig9"  
4699. .5833  
/note="assembly\_name:Contig10"  
5934. .8646  
/note="assembly\_name:Contig11"

misc_feature	8747. .11792	/note="assembly_name:Contig12"
misc_feature	11893. .14813	
misc_feature	14914. .19837	/note="assembly_name:Contig13"
misc_feature	19938. .24772	/note="assembly_name:Contig14"
misc_feature	24873. .30965	/note="assembly_name:Contig15"
misc_feature	31066. .40841	/note="assembly_name:Contig16"
misc_feature	40942. .52869	/note="assembly_name:Contig17"
misc_feature	52970. .63635	/note="assembly_name:Contig18"
misc_feature	63736. .90222	/note="assembly_name:Contig19"
misc_feature	clone_end:T7	
misc_feature	vector_side:right	
misc_feature	90323. .135921	/note="assembly_name:Contig21"
misc_feature	136022. .186935	/note="assembly_name:Contig22"
misc_feature	clone_end:SP6	
misc_feature	vector_side:left	
BASE COUNT 53462 a 40007 c 40377 g 51585 t 1504 others		
ORIGIN		
Query Match 6.8%; Score 95.4; DB 2: Length 186935;		
Best Local Similarity 44.0%; Pred. No. 1.3e-07;		
Matches 550; Conservative 0; Mismatches 691; Indels 10; Gaps 3;		
Qy	159	TAATAATGCTGCTGCTTTGCTTCCTGGTGTAGTACGCTGACACCTGTCACATAAATAA 218
Db	8430	TAATGCTGTTGCTGCTACTATTACTTGTGCTATATGCTGCTGCTGCTGCTGCTGCTG 8371
Qy	219	AGATGCTGCTGCTTAACCAATCCACCCTGCTACTGCTAAATTTAGTCACATAATGTAAG 278
Db	8370	TGCTACTGTTGCTGCTATGCTAAATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8311
Qy	279	TAATGCCCTGCTGGTACCGCAATTCGAGGTGGAGCAACAGATATGTCAGCAATATACAC 338
Db	8310	GATTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8251
Qy	339	AGAATGCTTAATGTPAGAAATTAATTTTATATGAATGCTGCTGCTGCTGCTGCTGCTG 398
Db	8250	TGCTTTTGTGCTATTGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8191
Qy	399	TGCTAGTACATGCACAGCTGTCGGTAAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 458
Db	8190	TGCTGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8131
Qy	459	TGCCGCTACCATAGTCCGCAATATGTAACGTGCGCATGCTGCTGCTGCTGCTGCTGCTG 518
Db	8130	TGCTGTTGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8071
Qy	519	TGGAGTAACCTACTGATTATGTTAGATCATTCACAGAAATGTTAAATGTAGACTTAAC 578
Db	8070	TGCTGATGCTGTTGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8011
Qy	579	TTACTATAATGTTAAATGTTAAATGTTAAATGTTAAATGTTAAATGTTAAATGTTAAAT 638
Db	8010	TGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7951
Qy	639	TTCGCGGCAATTTAAACCTGCTAAATGTTGCTTAAAGTACTTTTAGGTAATGATGCTACA 698
Db	7950	TGCTGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7891
Qy	699	AACCGCATATGTAACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 758
Db	7890	TATTGCTGCTACTATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7831







```

Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenwo, S., Oguh, M., Okwundu, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L. L., Quiles, M., Ren, Y.,
Rives, M., Rojas, A., Rojebokan, I., Rolfe, M., Ruiz, S., Savary, G.,
Scherer, S., Scott, G., Shen, H., Shooshtari, N., Sisson, I.,
Sodergren, E., Sonaike, F., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tatney, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 176822)
Worley, K.C.
Direct Submission
Submitted (09-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 176822)
Worley, K.C.
Direct Submission
Submitted (12-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 11, 2002 this sequence version replaced gi:18846108.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GCZV
Center clone name: CH230-98I1
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 115453 bases at least Q40
Consensus quality: 125501 bases at least Q30
Consensus quality: 133023 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 77 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1
* 1053: contig of 1052 bp in length
* 1152: gap of unknown length
* 1153: contig of 1027 bp in length
* 2179: gap of unknown length
* 2180: contig of 1076 bp in length
* 2280: gap of unknown length
* 3355: gap of unknown length
* 3456: contig of 1379 bp in length
* 4834: gap of unknown length
* 4935: contig of 1288 bp in length
* 6222: gap of unknown length
* 6223: contig of 1078 bp in length
* 6323: gap of unknown length
* 7401: contig of 1307 bp in length
* 7500: gap of unknown length
* 8808: contig of 1359 bp in length
* 8907: gap of unknown length
* 10266: contig of 1296 bp in length
* 10267: gap of unknown length
* 11662: contig of 1296 bp in length
* 11663: gap of unknown length
*
11763: contig of 1050 bp in length
12813: gap of unknown length
12912: contig of 1224 bp in length
14136: gap of unknown length
14236: contig of 1210 bp in length
15446: gap of unknown length
15447: contig of 1320 bp in length
16866: gap of unknown length
16966: contig of 1734 bp in length
18700: gap of unknown length
18800: contig of 1558 bp in length
20358: gap of unknown length
20458: contig of 1567 bp in length
22025: gap of unknown length
22125: contig of 2390 bp in length
24515: gap of unknown length
24615: contig of 1231 bp in length
25846: gap of unknown length
25946: contig of 1230 bp in length
27176: gap of unknown length
27276: contig of 1007 bp in length
28283: gap of unknown length
28383: gap of unknown length
29995: contig of 1612 bp in length
30095: gap of unknown length
31459: contig of 1364 bp in length
31559: gap of unknown length
32924: contig of 1365 bp in length
33024: gap of unknown length
34073: contig of 1049 bp in length
34173: gap of unknown length
35425: contig of 1252 bp in length
35525: gap of unknown length
36980: contig of 1455 bp in length
37080: gap of unknown length
38628: contig of 1548 bp in length
38728: gap of unknown length
40212: contig of 1484 bp in length
40312: gap of unknown length
42052: contig of 1740 bp in length
42152: gap of unknown length
43337: contig of 1185 bp in length
43437: gap of unknown length
45266: contig of 2089 bp in length
45626: gap of unknown length
46889: contig of 1263 bp in length
46890: gap of unknown length
48670: contig of 1681 bp in length
48770: gap of unknown length
50070: contig of 1300 bp in length
50170: gap of unknown length
51389: contig of 1219 bp in length
51489: gap of unknown length
53145: contig of 1656 bp in length
53245: gap of unknown length
53426: contig of 1375 bp in length
54620: gap of unknown length
54720: contig of 2200 bp in length
56920: gap of unknown length
57020: contig of 2045 bp in length
59065: gap of unknown length
59165: contig of 1547 bp in length
60712: gap of unknown length
60812: contig of 1064 bp in length
61876: gap of unknown length
61976: contig of 2802 bp in length
64778: gap of unknown length
64878: contig of 1914 bp in length
66792: gap of unknown length
66892: contig of 1683 bp in length
68575: gap of unknown length
68675: gap of unknown length
70134: contig of 1459 bp in length
70135: gap of unknown length
72222: contig of 1988 bp in length
70235

```









```
Center clone name: CH230-123D20
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 73854 bases at least Q40
Consensus quality: 79579 bases at least Q30
Consensus quality: 83355 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
      (see http://www.hgsc.bcm.tmc.edu/docs/Genbank\_draft\_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 57 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1
* 1478: contig of 1478 bp in length
* 1479: gap of unknown length
* 1579: contig of 1339 bp in length
* 2918: gap of unknown length
* 3017: gap of unknown length
* 4096: contig of 1079 bp in length
* 4196: gap of unknown length
* 5642: contig of 1446 bp in length
* 5742: gap of unknown length
* 6939: contig of 1197 bp in length
* 7039: gap of unknown length
* 8481: contig of 1442 bp in length
* 8581: gap of unknown length
* 10096: contig of 1515 bp in length
* 10196: gap of unknown length
* 11452: contig of 1256 bp in length
* 11453: gap of unknown length
* 12648: contig of 1096 bp in length
* 12649: gap of unknown length
* 14133: contig of 1385 bp in length
* 14233: gap of unknown length
* 15531: contig of 1318 bp in length
* 15651: gap of unknown length
* 17024: contig of 1373 bp in length
* 17124: gap of unknown length
* 18147: contig of 1023 bp in length
* 18247: gap of unknown length
* 19923: contig of 1676 bp in length
* 20023: gap of unknown length
* 21045: contig of 1022 bp in length
* 21145: gap of unknown length
* 22292: contig of 1147 bp in length
* 23292: gap of unknown length
* 23545: contig of 1153 bp in length
* 23645: gap of unknown length
* 25288: contig of 1643 bp in length
* 25388: gap of unknown length
* 26988: contig of 1600 bp in length
* 27088: gap of unknown length
* 29136: contig of 2048 bp in length
* 29236: gap of unknown length
* 30361: contig of 1125 bp in length
* 30461: gap of unknown length
* 31728: contig of 1267 bp in length
* 31828: gap of unknown length
* 33326: contig of 1498 bp in length
* 33426: gap of unknown length
* 34972: contig of 1546 bp in length
* 35072: gap of unknown length
* 37098: contig of 2026 bp in length
* 37198: gap of unknown length
* 38776: contig of 1678 bp in length
* 38976: gap of unknown length
* 40520: contig of 1544 bp in length
* 40620: gap of unknown length
*
* 1478: contig of 1478 bp in length
* 1479: gap of unknown length
* 1579: contig of 1339 bp in length
* 2918: gap of unknown length
* 3017: gap of unknown length
* 4096: contig of 1079 bp in length
* 4196: gap of unknown length
* 5642: contig of 1446 bp in length
* 5742: gap of unknown length
* 6939: contig of 1197 bp in length
* 7039: gap of unknown length
* 8481: contig of 1442 bp in length
* 8581: gap of unknown length
* 10096: contig of 1515 bp in length
* 10196: gap of unknown length
* 11452: contig of 1256 bp in length
* 11453: gap of unknown length
* 12648: contig of 1096 bp in length
* 12649: gap of unknown length
* 14133: contig of 1385 bp in length
* 14233: gap of unknown length
* 15531: contig of 1318 bp in length
* 15651: gap of unknown length
* 17024: contig of 1373 bp in length
* 17124: gap of unknown length
* 18147: contig of 1023 bp in length
* 18247: gap of unknown length
* 19923: contig of 1676 bp in length
* 20023: gap of unknown length
* 21045: contig of 1022 bp in length
* 21145: gap of unknown length
* 22292: contig of 1147 bp in length
* 23292: gap of unknown length
* 23545: contig of 1153 bp in length
* 23645: gap of unknown length
* 25288: contig of 1643 bp in length
* 25388: gap of unknown length
* 26988: contig of 1600 bp in length
* 27088: gap of unknown length
* 29136: contig of 2048 bp in length
* 29236: gap of unknown length
* 30361: contig of 1125 bp in length
* 30461: gap of unknown length
* 31728: contig of 1267 bp in length
* 31828: gap of unknown length
* 33326: contig of 1498 bp in length
* 33426: gap of unknown length
* 34972: contig of 1546 bp in length
* 35072: gap of unknown length
* 37098: contig of 2026 bp in length
* 37198: gap of unknown length
* 38776: contig of 1678 bp in length
* 38976: gap of unknown length
* 40520: contig of 1544 bp in length
* 40620: gap of unknown length
```

```
* 40621 41712: contig of 1092 bp in length
* 41713 41812: gap of unknown length
* 41813 43711: contig of 1899 bp in length
* 43712 43811: gap of unknown length
* 43812 45722: contig of 1911 bp in length
* 45723 45822: gap of unknown length
* 45823 47646: contig of 1824 bp in length
* 47647 47746: gap of unknown length
* 47747 49427: contig of 1681 bp in length
* 49428 49527: gap of unknown length
* 49528 51247: contig of 1720 bp in length
* 51248 51347: gap of unknown length
* 51348 53372: contig of 2025 bp in length
* 53373 53472: gap of unknown length
* 53473 55965: contig of 2493 bp in length
* 55966 56065: gap of unknown length
* 56066 58604: contig of 2539 bp in length
* 58605 58704: gap of unknown length
* 58705 61433: contig of 2729 bp in length
* 61434 61533: gap of unknown length
* 61534 63097: contig of 1564 bp in length
* 63098 63197: gap of unknown length
* 63198 65272: contig of 2075 bp in length
* 65273 65372: gap of unknown length
* 65373 67765: contig of 2393 bp in length
* 67766 67865: gap of unknown length
* 67866 68952: contig of 1087 bp in length
* 68953 72102: contig of 3050 bp in length
* 72103 72202: gap of unknown length
* 72203 74585: contig of 2383 bp in length
* 74586 74686: gap of unknown length
* 74687 77412: contig of 2727 bp in length
* 77413 77512: gap of unknown length
* 77513 80330: contig of 2818 bp in length
* 80331 80430: gap of unknown length
* 80431 82126: contig of 1696 bp in length
* 82127 82226: gap of unknown length
* 82227 84702: contig of 2476 bp in length
* 84703 84802: gap of unknown length
* 84803 88989: contig of 4187 bp in length
* 88990 89089: gap of unknown length
* 89090 92993: contig of 3904 bp in length
* 92994 93093: gap of unknown length
* 93094 96037: contig of 2944 bp in length
* 96038 96137: gap of unknown length
* 96138 100350: contig of 4213 bp in length
* 100351 100450: gap of unknown length
* 100451 104033: contig of 3583 bp in length
* 104034 104133: gap of unknown length
* 104134 106965: contig of 2832 bp in length
* 106966 107065: gap of unknown length
* 107066 112233: contig of 5168 bp in length
*
* Query Match 6.1%; Score 85.4; DB 2; Length 127354;
* Best Local Similarity 45.4%; Pred. No. 8.1e-06;
* Matches 308; Conservative 0; Mismatches 371; Indels 0; Gaps 0;
*
* Qy 614 ATCCAGGTAAAGTTAATGACACACCTTGTCGGGAATTAACCTGCTAATGTTGCTTAAG 673
* Db 118142 AGCAAGACTAAACTGAATGGTTACTACTGCTACTACTACTACTGCTGCTACTACTACAG 118201
*
* Qy 674 CTACTTTAGGTAATGCTACATAACCCGCAATGTAACGTTGATGCTGCTGCTA 733
* Db 118202 CTGCTATTACTACTACTGCTACTACTACTACTACTACTACTACTACTACTACTACTGCTA 118261
*
* Qy 734 CTATAAGTGTGCTGGAGTAATAATTTGGGTAGCAGCAAAACACTGAATGTACTAATTTG 793
* Db 118262 CTACTGCTGCTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 118321
*
* Qy 794 CTCCTTAACCTTTTACAAATAAATGCTCCTTAATTCATCCAGGTAATAGTACATGCTAC 853
* Db 118322 CTGCTGCTACTACTACTGCTGCTACTACTACTACTGCTGCTACTACTACTACTACTAATA 118381
```



[illegible]



